DOCUMENT RESUME

ED 471 306 TM 034 675

| | |
|---|---|
| AUTHOR | Daniel, Larry G.; Onwuegbuzie, Anthony J. |
| TITLE | Reliability and Qualitative Data: Are Psychometric Concepts Relevant within an Interpretivist Research Paradigm? |
| PUB DATE | 2002-11-00 |
| NOTE | 22p.; Paper presented at the Annual Meeting of the Mid-South Educational Research Association (Chattanooga, TN, November 6-8, 2002). |
| PUB TYPE | Reports - Descriptive (141) -- Speeches/Meeting Papers (150) |
| EDRS PRICE | EDRS Price MF01/PC01 Plus Postage. |
| DESCRIPTORS | Models; *Qualitative Research; *Reliability |
| IDENTIFIERS | *Interpretivism; *Positivism |

ABSTRACT

Reliability is one of the chief characteristics researchers consider when judging the quality of data used in their studies. Within the positivist paradigm, data are typically quantified, and thus it is relatively easy to derive estimates of reliability. Within the interpretivist paradigm, however, the idea of data reliability is a looser science. This paper makes the case that the positivist and interpretivist paradigms are not as disparate as many suppose in terms of conceptualizations of reliability. A variety of methods for assessing the reliability, or trustworthiness, of qualitative data are reviewed, including the important process of triangulation. Terminology appropriate to specific data features that affect reliability is compared across the positivist and interpretivist paradigms. (Contains 1 table and 31 references.) (Author/SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 471 306



Running Head: RELIABILITY AND QUALITATIVE DATA

# Reliability and Qualitative Data: Are Psychometric Concepts Relevant Within an Interpretivist Research Paradigm?

Larry G. Daniel

University of North Florida

Anthony J. Onwuegbuzie

Howard University

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

L. G. Daniel

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

TM034675

Paper presented at the annual meeting of the Mid-South Educational Research Association, Chattanooga, TN, November 6-8, 2002. Please direct all correspondence to the first author at College of Education and Human Services, 4567 St. Johns Bluff Road South, Jacksonville, FL 32224-2676 (email: ldaniel@unf.edu).

BEST COPY AVAILABLE

Abstract

Reliability is one of the chief characteristics researchers consider when judging the quality of data utilized in their studies. Within the positivist paradigm, data are typically quantified, and, thus, it is relatively easy to derive estimates of reliability. Within the interpretivist paradigm, however, the idea of data reliability is a looser science. In the present paper, we argue that the positivist and interpretivist paradigms are not as disparate as many suppose in terms of conceptualizations of reliability. A variety of methods for assessing the reliability, or trustworthiness, of qualitative data are reviewed, including the important process of triangulation. Terminology appropriate to specific data features that affect reliability are compared across is compared across the positivist and interpretivist paradigms.

# Reliability and Qualitative Data: Are Psychometric Concepts Relevant Within an Interpretivist Research Paradigm?

Reliability is one of the chief characteristics researchers consider when judging the quality of data utilized in their studies. Within the positivist paradigm, data are typically quantified, and, thus, it is relatively easy to derive estimates of reliability based on various statistical indices developed for this purpose (Pedhazur & Schmelkin, 1991). In qualitative research, however, the idea of data reliability is a looser science, considering that the researcher serves as the instrument and that the researcher's understandings and interpretations serve as the data gathered with the "instrument." Consequently, some have argued that reliability of qualitative findings cannot (and should not) be estimated or assessed at all. In fact, many who advocate for the importance of an interpretivist research paradigm (e.g., Smith, 1984) refrain from using the term "reliability," fearing that the positivist framework of reliability will be considered as the standard against which all data integrity issues are conceptualized and assessed.

In the present paper, however, we argue that the positivist and interpretivist paradigms are not as disparate as many suppose in terms of conceptualizations of issues surrounding reliability. Logical connections between the two paradigms as regards reliability issues are discussed, and a list of terminology is presented to illustrate how 12 specific data features relative to reliability are addressed within the two paradigms.

## Reliability Within the Interpretivist Paradigm

In qualitative research, information gleaned from observations, interviews, and the like must be "trustworthy" (Eisenhart & Howe, 1992; Lincoln & Guba, 1985); otherwise any themes that emerge from these data will not be credible. An important component of trustworthiness is "dependability" (Lincoln & Guba, 1985). Interestingly, dependability is analogous to reliability (Eisenhart & Howe, 1992; Onwuegbuzie, in press), and, the term was perhaps used originally by Cronbach, Gleser, Nanda, and Rajaratnam (1972) to refer to a rather classical/postivist view of reliability vis-a-vis generalizability theory. Onwuegbuzie (in press) identified 24 methods for assessing the trustworthiness of qualitative data. Many of these techniques can be utilized to assess the dependability or reliability of qualitative data extracted. Techniques for evaluating this dimension of trustworthiness include triangulation, which involves the use of multiple and different methods, investigators, sources, and theories to obtain corroborating evidence (Ely, Anzul, Friedman, Garner, & Steinmetz, 1991; Glesne & Peshkin, 1992; Lincoln & Guba, 1985; Merriam, 1988; Miles & Huberman, 1984, 1994; Onwuegbuzie, in press; Patton, 1990).

Triangulation reduces the possibility of chance associations, as well as of systematic biases prevailing due to a specific method being utilized, thereby allowing greater confidence in any interpretations made (Fielding & Fielding, 1986; Maxwell, 1992). Hence, Lancy (1993, p. 20) noted, "The qualitative researcher's most effective defense against the charge of being subjective is to buttress what she has observed

with material that reinforces these observations from other semi-independent sources." Likewise, Eisner (1998) proposed "structural corroboration" as a synonym for triangulation, noting, "Structural corroboration is the term I use to describe the confluence of multiple sources of evidence or the recurrence of instances to support a conclusion" (p. 55).

According to Denzin (1978), three outcomes arise from triangulation: convergence, inconsistency, and contradiction. Each of these outcomes clearly represents issues pertaining to reliability. Nevertheless, many interpretivists refrain from using the term "reliability" when pertaining to qualitative data, probably because of an attempt to distance qualitative analytical techniques from statistical method (Madill et al., 2000). However, this line of thinking is counterproductive. Indeed, as noted by Constas (1992, p. 255), unless methods for examining rival hypotheses in qualitative research are developed, "the research community will be entitled to question the analytical rigor of qualitative research"--where rigor is defined as the attempt to make data and categorical schemes as public and as replicable as possible (Denzin, 1978).

## Analyzing and Comparing Reliability Issues Across Paradigms

As previously noted, we maintain that issues relative to reliability of social science data do not vary appreciably across the positivist and interpretivist paradigms, with specific data features that affect reliability being constant across the paradigms. The major differences revolve around the nature of the data and the philosophical assumptions of the paradigms. Hence, terminology has developed that is distinctive to

each paradigm. As presented in Table 1, we have identified 12 key features of data that affect reliability. Each of these features is discussed in light of its applicability to the two paradigms.

*Consistency of Evidence*

Within the positivist paradigm, response variance on variables of interest is the focus of most all data analyses. In a conceptual sense, reliability coefficients are an estimate of the percent of the total variance in the scores on the measurement of interest that is attributable to true score variance (Cronbach, 1951). When this estimate is high, the researcher has enough evidence to place confidence in the scores and in the scores' use in additional descriptive, parametric, or non-parametric analyses.

Within the interpretivist paradigm, consistency of evidence is defined more loosely as the degree to which the data are "trustworthy." Although the term "trustworthiness" is defined in varying ways, it seems generally to cover at least some of the issues addressed by "research validity," "measurement validity" and "measurement reliability" within the positivist paradigm. For example, Lincoln and Guba (1985) posed four standards that should be used when judging qualitative, or naturalistic, studies: credibility, transferability, and confirmability, and dependability (or consistency). Credibility and transferability would apply to both research validity and measurement validity, whereas dependability (consistency) would be a standard for judging something roughly equivalent to reliability. Confirmability (i.e., objectivity) would be applicable across all areas. Wolcott (1990) noted that consistency is the degree to

which a study is free of inner contradictions, cautioning, however, against researchers assuming a totally contradiction-free approach as this would "set us to wondering how they could be [accurately] describing human behavior" (p. 134). Trustworthiness, therefore would involve looking for a high degree of consistency in the findings and presenting an explanation for factors to which any inconsistent findings might be attributed.

*Data Integrity*

Within the postivist paradigm (and more particularly within the circles of classical measurement theorists), it is commonplace for researchers to speak of the "psychometric integrity of the data," a term that normally implies some set of assumptions about validity, reliability, and other related measurement characteristics (Crocker & Algina, 1986). One would normally expect to see at least one estimate of reliability among whatever other data might accompany the reference to data integrity. For the interpretivist, data integrity would be essentially equivalent to the "consistency of evidence" and would refer to consistency or dependability of the data. Dependability is often addressed in terms of data triangulation, with a variety of qualitative data collection and analysis strategies used simultaneously and, in some cases, supplemented with quantitative methods in a mixed methods approach.

*Consistency of Judgments or Interpretations*

Some measures of performance-based tasks in education and related disciplines (e.g., writing samples, public speaking, teaching behaviors) require the rater

or scorer to exercise a moderate to appreciable amount of judgment when determining scores for the individual or work sample being rated. Consequently, psychometric specialists have developed various indices of inter-rater agreement and inter-rater reliability. These may be in the form of correlation-type indices or degree-of-difference indices. Calculation of these indices allows for (a) tracking of the consistency (i.e., fairness) of the scoring process across raters and (b) gaining evidence to substantiate possible rater effects that might contaminate the scores (e.g., rater severity, biases, inconsistency in application of scoring criteria).

Fortunately, rater agreement as a concept in qualitative data analysis is increasingly gaining acceptance. In particular, it is no longer unusual for qualitative researchers to report either intrarater (i.e., consistency of a given rater's scores or observations–in essence, a variation of test-retest reliability) or interrater (i.e., consistency of two or more independent raters' scores or observations) reliability estimates (Gay & Airasian, 2000). Evidence of rater agreement can be gleaned from the fact that a leading theory-building qualitative software program called NUD.IST (non-numerical unstructured data indexing searching & theorizing) allows data analysts to determine inter-coder reliability (QSR International Pty Ltd., 2002). Even in the absence of these inter-coder issues, however, it is important for qualitative researchers to realize that all data, regardless of their nature or how they are collected, are subject to the limitations of the specific conditions under which they have been collected (Marshall & Rossman, 1999).

*Temporality of the Data*

Positivist researchers should be aware of the temporal nature of data and the degree to which temporality can affect reliability and correlation. Quantitative data within the social sciences are subject to conditions of "temporal instability," namely the tendency for scores on variables of interest to fluctuate over relatively short periods of time. Granted, scores on certain measures would be expected to change over longer time periods due to maturity, effectiveness of interventions, or other natural or imposed changes that take place within an individual over a reasonable period of time. In other cases, scores will tend to vary without a reasonable explanation within a relatively brief period of time, making the data suspect due to temporal instability. Further, as Nunnally (1994) noted, "a measure which has low temporal stability will not be a good predictor of future behavior" (p. 243). Within the interpretivist paradigm, temporality is played out in terms of the relativism, or context specificity, of the data. Bernstein (1983) noted that any reality under study "must be understood as relative to a specific conceptual scheme, theoretical framework, paradigm, form of life, society or culture" (p. 685).

*Corroboration of Evidence from Multiple Sources*

In traditional measurement integrity studies, coefficients of equivalence are used in cases in which multiple forms of a test have been developed. Participants would be administered both forms of the test, and correlations between the two forms would be computed. Higher coefficients would imply that data from one test are equally meaningful as the other, providing evidence that the construct of interest can be

measured as effectively with one score as with the other. Further, if subjects are tested across a variety of scenarios that simultaneously explore several facets of measurement (e.g., internal consistency, occasion of measurement, and equivalent forms), generalizability theory analysis could be used to better understand the effects of those multiple sources of variation in test scores.

As previously noted, within the interprevist paradigm, researchers may utilize triangulation procedures (i e., structural corroboration) to achieve a similar result as obtained when testing for equivalence within the positivist paradigm. Originally a term used in navigation and surveying, triangulation serves as a method for the qualitative analyst to "steer the course" in the direction of a more accurate data interpretation. This gives the researcher an opportunity to account for the strengths and weaknesses of each data collection strategy and to examine the overall data for convergence toward a clear understanding of the phenomena under consideration: "Triangulation assumes that looking at an object from more than one standpoint provides researchers and theorists with more comprehensive knowledge about the object" (Miller, 1997, p. 25).

*Cohesiveness of Evidence*

Because experimentation is prohibitive in many practical measurement situations (e.g., it is difficult to do the test-retest studies needed to assess for score stability within a regular first-grade classroom), researchers often limit themselves to reliability studies that feature internal consistency measures (e.g., Cronbach alpha [Cronbach, 1951], K-R 20 and K-R 21 estimates [Kuder & Richardson, 1937]). Scores

gathered via a common set of items can yield first line evidence of reliability. Likewise, a series of related pieces of qualitative data that are used to form a narrative argument can be examined collectively for evidence of "coherence." Eisner (1998) contended that coherence is rooted in "believability" of qualitative findings and the law of good fit, noting, "We scrutinize the argument by looking for inconsistencies, lapses of logic, things that just don't fit" (p. 53).

*Data Inconsistency*

The notion of cohesiveness leads logically to an antithetical concept, namely, data inconsistency. Within the positivist paradigm, reliability is diminished when scores contain higher amounts of measurement error, or unexplained/unsystematic variance. Problems with error can be tracked and reported using various estimates of standard error. Obviously, the interpretivist does not normally have tight quantitative data available to make these types of judgments; however, it is possible through triangulation for the interpretivist to explore inconsistencies in the findings and, in larger data sets, to look for negative cases (i.e., cases that stand out as atypical in terms of the relationships among the phenomena of interest in the given qualitative study). These inconsistencies can be useful in generating theories for investigation in future studies (Woods, 1992).

*Alternate Explanations*

Traditional reliability analyses can sometimes yield totally unexpected and seemingly illogical results. For example, a reliability coefficient can be negative,

indicating, at least *prima facie*, that less than zero percent of the variance is being attributed to the true score! Obviously, such a finding is logically impossible even though it is mathematically accurate. Erroneous values should cue the researcher to examine the data further for evidence of rival hypotheses that account for the data better than the researcher's original hypothesis. In the present example, there may be two competing constructs underlying the data. The researcher might wisely look for evidence to support a rival hypothesis that the items, upon being split into two sets, might yield more reasonable coefficients. Similarly, Denzin (1978), noted that "contradictions" are a likely outcome of triangulation procedures used in qualitative case studies. Similar to inconsistencies, contradictions represent broader patterns within the data in which data from one source do not "line up" with data from other sources. Contradictions may indicate a systematic misunderstanding of the data, a larger concern regarding the reliability of the data, or the need to develop a new theory to support the data if the new theory is determined to be a legitimate representation of the reality being studied.

*Surety of Evidence*

Estimates of standard error provide the positivist with a means for assessing the surety (or accuracy) of a given data result. Standard errors can be utilized to develop confidence intervals around reliability coefficients or descriptive statistics generated for a variable of interest. If the standard error is low, resulting in a small confidence interval, the researcher can place confidence in the result.

Within the interpretivist paradigm, confidence in (surety of) in the results has a lot more to do with how well that data have been recorded and coded. Data coding allows for the assignment of alpha-numeric symbols to various observations for purposes of tracking the incidence of phenomena of interest. Obviously, these codes would have to be applied uniformly and consistently if any confidence or surety was to be placed in the data generated by the coding processes. Hence, Kelle and Laurie (1995, pp. 24-25) noted:

> a coding frame[work] would only be regarded as reliable if in any subsequent re-coding exercise the same codes could be applied to the same incidents, which means that the coding could be repeated by a different coder within an acceptable margin of error. To attain this goal one would be careful to construct coding categories which are mutually exclusive and unambiguous. . . . [I]t is of crucial importance to apply these codes consistently to ensure that the same text segments are assigned the same codes, since otherwise different members of the same research group would draw upon different information when referring to the same topics.

*Elusive Goal of Data Collection*

Within any area of inquiry is embedded some ultimate goal or reality which the researcher hopes to attain. While these goals are typically elusive, if not utopian, it is important that researchers keep these goals in mind when examining actual results that are obtained from a given study within that area of inquiry. Classical measurement

theorists who conduct reliability analyses are ultimately searching for an accurate understanding of the true score for each individual in a given sample. The true score may be defined as "the average score that would be obtained over repeated testings" (Nunnally, 1994, p. 211). Of course, to actually obtain the true score, the number of "repeated testings" needed would be prohibitive, not to mention that each test score gained via these repetitions would be contaminated by error such that the true score would still remain elusive.

The equivalent elusive goal within the interpretivist paradigm is the capturing of the social understanding or social reality underlying the events, activities, and behaviors being studied. Roman (1992) distinguished between the "practices behaviors, and social meanings arising in the field when a researcher is physically present among the research subjects and when she or he is physically absent" (p. 571). These two social realities are clearly distinct, and even if the researcher argues for the former reality, the elusivity issue still exists considering that reality changes moment by moment, resulting in "the impossibility of knowing the world in its pristine state" (Eisner, 1998, p. 46).

*Data Collection Setting*

All data collected in any study are subject to the limitations of the scenario in which they were collected. For example if a teacher were to administer a test to a group of students, a host of factors related to the occasion on which the test was given might have an impact on the results (e.g., individual differences in the achievement levels of

the students, degree of fatigue or anxiety a given student is experiencing, inability of a student in understanding the test directions). This entire set of factors that makes the collection of data on any measure somewhat unique is what is referred to in classical measurement circles as the "occasion of measurement." Within the positivist paradigm, researchers postulate that these "temporal" factors are always present to some degree, and attempts are made to estimate the effects of these factors on the reliability of scores via the computation of coefficients of stability or other similar reliability indices.

As is true in quantitative studies, there is clearly an "occasion of measurement" for any data collected in a qualitative study. However, the nature of the qualitative data precludes the type of statistical analyses used in quantitative approaches when examining the impact of the specific scenario in which data were collected on the researcher's perception of the results. For example, a qualitative analyst may use a strictly narrative approach to cataloging observational data with no generation of performance scores or other scaled criteria with which the narrative data could be triangulated. In this case, the researcher's findings would be limited by this single observational setting, and the degree to which the findings would generalize, in absence of additional confirming evidence, to other similar settings would be unknown.

*Adequacy of Evidence*

A final feature of reliability, adequacy of evidence, has implications for both positivists and interpretivists. To the positivist, reliability evidence is subject to the adequacy of the researcher's "domain sampling." In preparing a measurement tool, the

researcher selects a sample of test items (or other data prompts) of some size from a given domain (population, universe) of all possible items. Even though scores on the measurement tool may be shown to be reliable, Nunnally (1994) warned that the amount of evidence one has for making a decision about the meaning is limited to the extent that the selected items are deemed not to adequately reflect the entire domain. This concern for adequacy of evidence is essentially a reminder that reliability does not equal validity.

Similarly, interpretivists must be concerned with the degree to which the narrative descriptions provide an adequate view of the social phenomena of interest. Within this paradigm, the "thickness" of the description will have an impact on the adequacy of the evidence (Marshall & Rossman, 1999). As Eisner (1998, p. 15) noted, "Thick description is an effort aimed at interpretation, at getting below the surface to that most enigmatic aspect of the human condition: the construction of meaning." If the researcher's description is overly superficial, the result will be data that are consistent (reliable) to some degree but that will fall short of the trustworthiness criterion expected of good qualitative research.

References

Bernstein, R. J. (1983). *Beyond objectivism and relativism: Science, hermeneutics, and praxis*. Philadelphia, PA: University of Pennsylvania Press.

Constas, M. A. (1992). Qualitative analysis as a public event: The documentation of category development procedures. *American Educational Research Journal, 29*, 253-266.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Orlando, FL: Holt, Rinehart, and Winston.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297-334.

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of educational measurements: Theories of generalizability for scores and profiles*. New York: Wiley.

Denzin, N. K. (1978). *The research act: A theoretical introduction to sociological methods*. New York: Praeger.

Eisenhart, M. A., & Howe K. R. (1992). Validity in educational research. In M.D. LeCompte, W.L. Millroy, & J. Preissle (Eds.), *The handbook of qualitative research in education* (pp. 643-680). San Diego, CA: Academic Press.

Eisner, E. W. (1998). *The enlightened eye: Qualitative inquiry and the enhancement of educational practice*. Upper Saddle River, NJ: Prentice-Hall.

Ely, M., Anzul, M., Friedman, T., Garner, D., & Steinmetz, A. C. (1991). *Doing qualitative research: Circles within circles*. New York: Falmer.

Fielding, N., & Fielding, J. (1986). *Linking data*. Beverly Hills, CA: Sage.

Gay, L. R., & Airasian, P. W. (2000). *Educational research: Competencies for analysis and application* (6th ed.). Englewood Cliffs, N.J.: Prentice Hall.

Glesne, C., & Peshkin, A. (1992). *Becoming qualitative researchers: An introduction*. White Plains, NY: Longman.

Kelle, U., & Laurie, H. (1995). Computer use in qualitative research and issues of validity. In U. Kelle (Ed.), *Computer-aided qualitative data analysis* (pp. 19-28). London: Sage.

Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika, 2*, 151-160.

Lancy, D. F. (1993). *Qualitative research in education: An introduction to the major traditions*. New York: Longman.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.

Marshall, C., & Rossman, G. B. (1999). *Designing qualitative research* (3rd ed.). Thousand Oaks, CA: Sage.

Maxwell, J. A. (1992). Understanding and validity in qualitative research. *Harvard Educational Review, 62*, 279-299.

Merriam, S. (1988). *Case study research in education: A qualitative approach*. San Francisco, CA: Jossey-Bass.

Miles, M. B., & Huberman, A. M. (1984). Drawing valid meaning from qualitative data: Toward a shared craft. *Educational Researcher, 13*, 20-30.

Miles, M., & Huberman, M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.

Miller, G. (1997). Building bridges: The possibility of analytic dialogue between ethnogrpahy, conversation analysis and Foucault. In D. Silverman, *Qualitative research: Theory, method and practice*. London: Sage.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Onwuegbuzie, A. J. (in press). Validity and qualitative research: An oxymoron? *Current Issues in Education*.

Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). Newbury Park, CA: Sage.

Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Lawrence Erlbaum.

QSR International Pty Ltd. (2002). *N6 (Non-numerical unstructured data indexing searching & theorizing) qualitative data analysis program*. (Version 6.0) [Computer software]. Melbourne, Australia: QSR International Pty Ltd.

Roman, L. G. (1992). The political significance of other ways of narrating ethnography: A feminist materialist approach. In M. D. LeCompte, W. L. Millroy, & J. Preissle,

*The handbook of qualitative research in education* (pp. 555-594). San Diego, CA: Academic Press.

Smith, J. K. (1984). The problem of criteria for judging interpretive research. *Education Evaluation and Policy Analysis, 6*, 379-391.

Wolcott, H. F. (1990). On seeking–and rejecting–validity in qualitative research. In E. W. Eisner & A. Peshkin, *Qualitative inquiry in education: The continuing debate* (pp. 121-152). New York: Teachers College Press.

Woods, P. (1992). Symbolic interactionism: Theory and method. In M. D. LeCompte, W. L. Millroy, & J. Preissle (Eds.), *The handbook of qualitative research in education* (pp. 337-404). San Diego, CA: Academic Press.

Table 1

Comparison of Terminology Relative to Reliability of Measurements/Data Across Two Research Paradigms

| DATA FEATURES | PARADIGM | |
|---|---|---|
| | **Positivist Terminology** | **Interpretivist Terminology** |
| consistency of evidence | explained variance | trustworthiness, consistency |
| data integrity | reliability, psychometric integrity | dependability, consistency |
| consistency of judgments or interpretations | inter-rater agreement | consensus |
| temporality of the data | temporal stability | relativism |
| corroboration of evidence from multiple sources | equivalency, generalizability | triangulation, structural corroboration, convergence |
| cohesiveness of evidence | internal consistency | coherence |
| data inconsistency | measurement error, unexplained variance | inconsistency, negative case analysis |
| alternate explanations | rival hypotheses | contradictions |
| surety of evidence | confidence interval | accuracy of coding |
| elusive goal of data collection | true score analysis | social understanding, social reality |
| data collection setting | occasion of measurement | observational setting |
| adequacy of evidence | domain sampling | thick description |

**U.S. Department of Education**
*Office of Educational Research and Improvement (OERI)*
*National Library of Education (NLE)*
*Educational Resources Information Center (ERIC)*

ERIC Educational Resources Information Center

# REPRODUCTION RELEASE

(Specific Document)

TM034675

## I. DOCUMENT IDENTIFICATION:

| | |
|---|---|
| Title: Reliability and Qualitative Data: Are Psychometric Concepts Relevant Within an Interpretivist Research Paradigms | |
| Author(s): Larry G. Daniel | |
| Corporate Source: Univ. of North Florida Jacksonville, FL 32224 | Publication Date: Nov. 2002 |

## II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Larry G. Daniel

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

Level 1

☒

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) and paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

Level 2A

☐

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2B

Level 2B

☐

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

*I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.*

**Sign here, → please**

| | |
|---|---|
| Signature: Larry G. Daniel | Printed Name/Position/Title: Larry G. Daniel |
| Organization/Address: Univ. of North Florida / COEHS Jacksonville, FL 32224-2676 | Telephone: 904 620 2522 FAX: 904 620 2522 |
| | E-Mail Address: lDaniel@unf.edu Date: 11/8/02 |

(Over)

## III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

| |
|---|
| Publisher/Distributor: |
| Address: |
| Price: |

## IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

| |
|---|
| Name: |
| Address: |

## V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC CLEARINGHOUSE ON ASSESSMENT AND EVALUATION
UNIVERSITY OF MARYLAND
1129 SHRIVER LAB
COLLEGE PARK, MD 20742-5701
ATTN: ACQUISITIONS

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

**ERIC Processing and Reference Facility**
**4483-A Forbes Boulevard**
**Lanham, Maryland 20706**

**Telephone:** 301-552-4200
**Toll Free:** 800-799-3742
**FAX:** 301-552-4700
**e-mail:** info@ericfac.piccard.csc.com
**WWW:** http://ericfacility.org

EFF-088 (Rev. 2/2001)